

FIG. 1

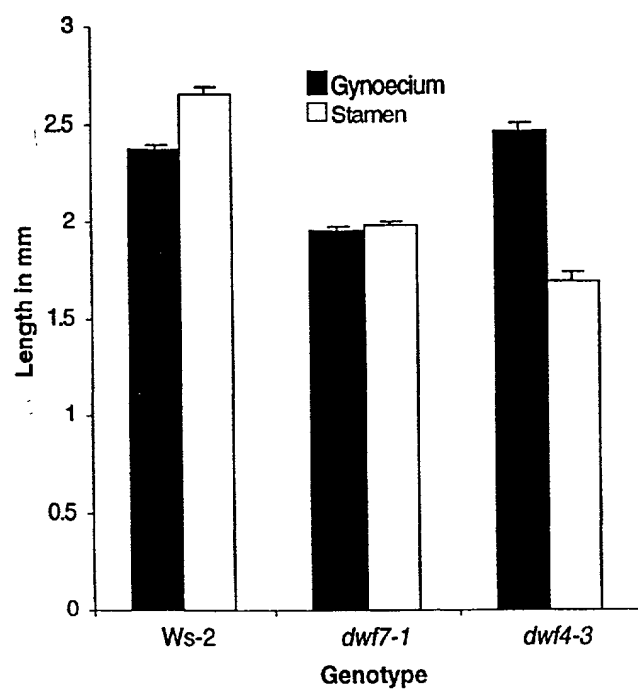


FIG. 2

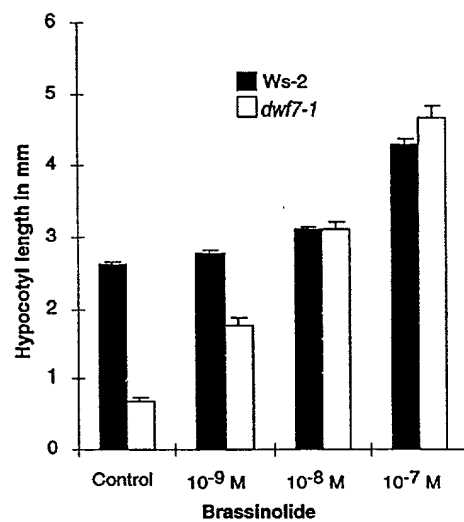


FIG. 3

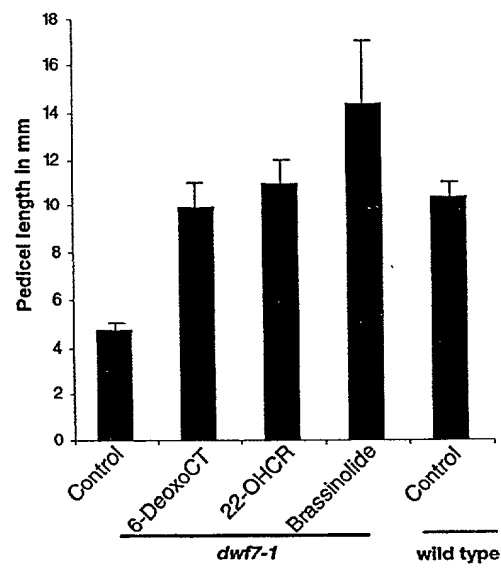


FIG. 4

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

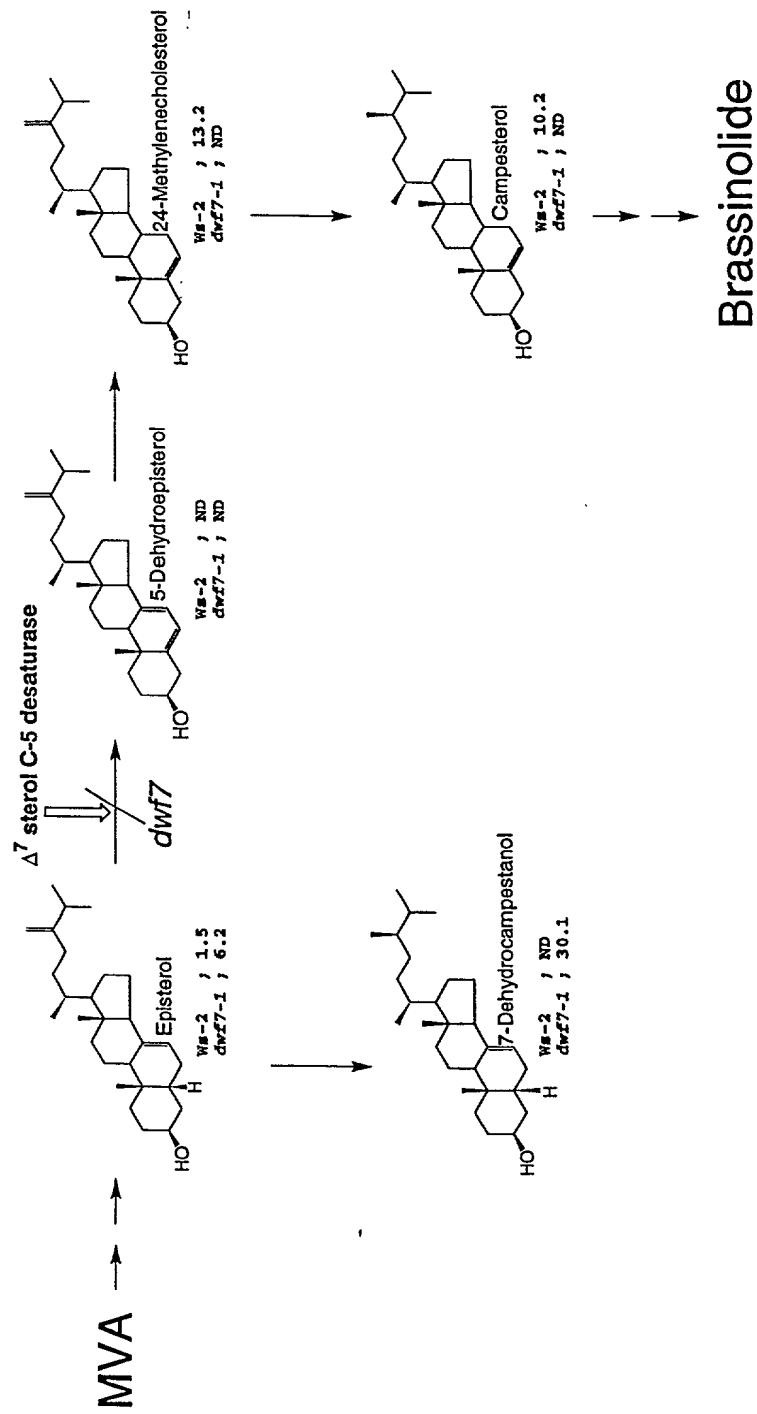


FIG. 5

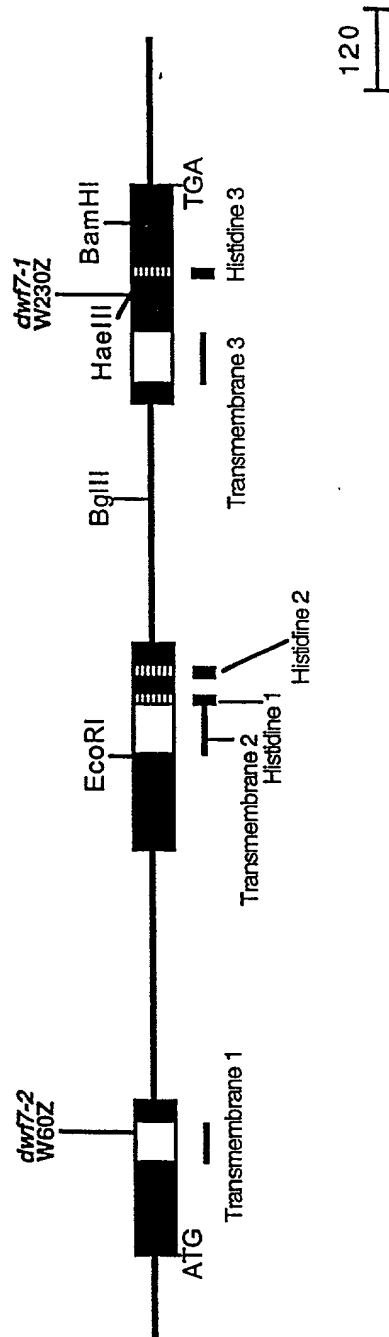


FIG. 6

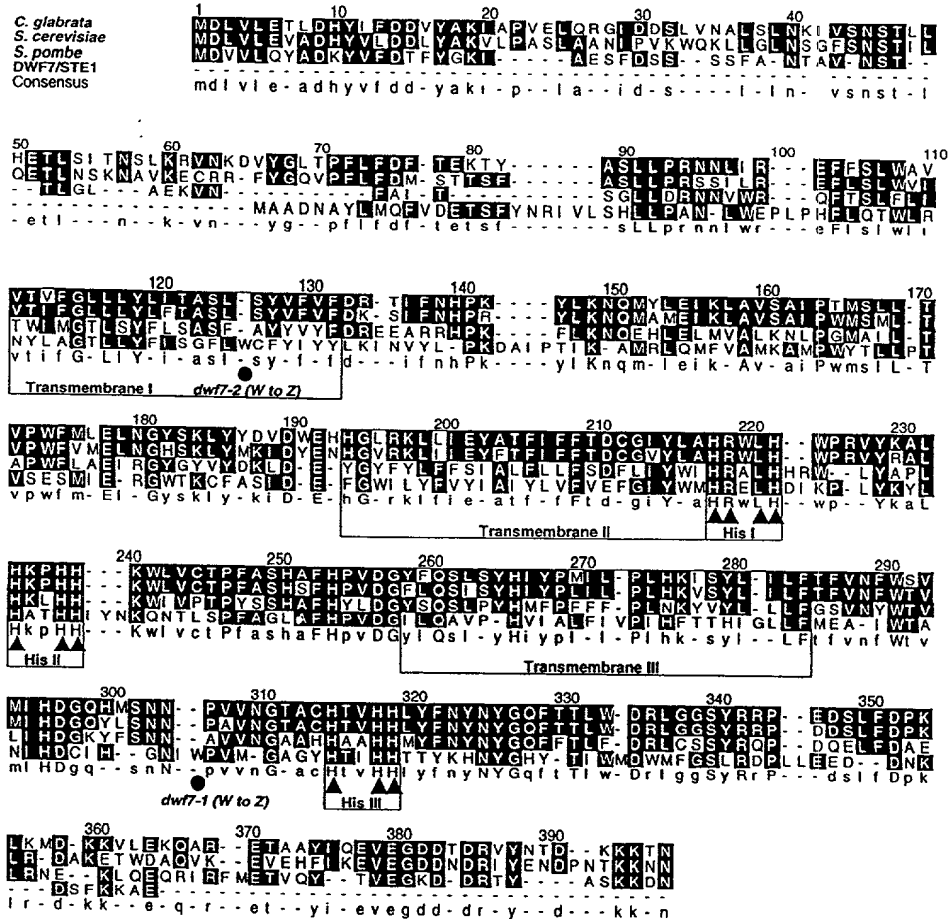


FIG. 7

10 20 30 40
 GAAGATCGATCAATCAATCATCAAACCTCTCTGTGTGCCAC

50 60 70 80 90 100
 41 ATGCATTACT ACTGTTGACT TGTTCAATTA AGGTAAGTA AGATCAATCC GCGCAATCTT

110 120 130 140 150 160
 101 CTTTCGTTTT CCGGCACCGA TCTCGGTGGA TCTCCGATTG ACATGCCGGC GGAATATGCT

170 180 190 200 210 220
 161 TATCTGATGC AGTTTGTTCG CGAATCTCT TTTTACACCC GAATCGTTCT GATCATCTT

230 240 250 260 270 280
 221 TTGCCCGCGA ATCATGCGGA ACCCTTACCT CATTTTCTCC AGACATGGCT CCGAATTAC

290 300 310 320 330 340
 281 CTCGCCGGAA CCTACTATA CTTCATCTCC GGTTCCTCT GGTGCTTCTA CATCTATTAC

L A G T L L Y F I S G F L W C F Y I Y Y

350 360 370 380 390 400
 341 CTTAAATCA AGGTTTACCT TCCCAAGGT CTCGACTTTC ACTTTGTAT TCATATTGC

410 420 430 440 450 460
 401 TTATCCGCTT TCTATCTTAT CGATTITTCG ATTTAGCGAA CAGGCTTCT TCCTGACTG

470 480 490 500 510 520
 461 TACACTAAT TCATTTGAT GTGGATAGT CATGTTTCCA TTATTCATT ATTGTGCTT

530 540 550 560 570 580
 521 ATTCTCCATC TAAGCGATTG AACATTTAGT GCCTTATATA ACTTTTGTG CACCCATGA

590 600 610 620 630 640
 581 GAATCGTAC ATCTTGAGG TTGATTTTC TACTGCCAT TTAGTCCAC TTAATTGTT

FIG. 8A

650 660 670 680 690 700
 641 TGTTCAGTC ATTCTCTACT TTCAGACACA TTTTITTTCT GCTTCTCTGA CACTCTGTCT
 710 720 730 740 750 760
 701 TAGTTTGAAA TCTTTTITGG TCTGTTTGGC TTCAGATGCA ATTCTACAA TAAAGGCTAT
 D A I P T I K A M
 770 780 790 800 810 820
 761 GCGTTGCAG ATGTTTGTCG CATTGACGC TATGCCATCG TACTCTTTC TTCCACTGT
 R L Q M F V A M K A M P W Y T L L P T V
 830 840 850 860 870 880
 821 CTCGACAGT ATGATTGACG GTGGTTGCAC CAAATGTTTT GCTAGCATAG ACCAATTCGG
 S E S M I E R G H T K C F A S I D E F G
 890 900 910 920 930 940
 881 CTGGATTCTG TATTTTGTTT ACATGCCCAT CTATCTTGTI TTCGTTGAGT TTGCTATTTA
 W I L Y F V Y I A I Y L V F V E F G I Y
 950 960 970 980 990 1000
 941 TTGATGCGAC ACAGACCTTC ATGACATTAA CCTCTCTAT AAGTATCTCC ATGCCACCCA
 H M H R E L H D I K P L Y K Y L H A T H
 1010 1020 1030 1040
 1001 TCATATCTAC AACAGGCAGA ATACACTCTC TCCATTGGCC G
 H I Y N K Q H T L S P F A

FIG. 8B

1050 1060 1070 1080
 GTAGTGTG TTCAGTTGT TCTCTTTAG TTCTGTAAG

1090 1100 1110 1120 1130 1140
 1081 AGATTGGTAG CATTTAGTTT CTTACCAGAA AAGACTTTGT CAGCAGCTGC TTGTACTCCA

1150 1160 1170 1180 1190 1200
 1141 AATCAGATTG TGCATTCCTT ATCCATAAGC TAACCAGAAA GGCTAGATTG ATATAATCT

1210 1220 1230 1240 1250 1260
 1201 CAGCTGCATT ACTTCACATA TGTACAGAGC ACTTCTGACT TAACCAGAGT TTGATCTTT

1270 1280 1290 1300 1310 1320
 1261 GTGTTTCTCT TCTGCTCTCG GACTGATTGG AATGACCGC AAGTCTTTT ATCTACTTC

1330 1340 1350 1360 1370 1380
 1321 CTGGAGTGTA TCTTGCTTAA TCCAGCGATG TGACATCTAA TATTACTTGT AACTTCCTTA

1390 1400 1410 1420 1430 1440
 1381 CTTTTTGTG TACAGCGCTT GCATTCACG CAGTAGACCG GATACTTCAG GCTGTACCGC

G L A F H P V D G I L Q R V P

1450 1460 1470 1480 1490 1500
 1441 ATGTCATAGC GCTGTTTATA GTGCCATTG ATTCACACG TCATATAGG CTTTGTCTA

H V I A L F I V P I H F T T H I G L L F

1510 1520 1530 1540 1550 1560
 1501 TGGAGCGGAT ATGGACCGCG AATCCATG ACTGCATCCA TGGCACATC TGGCCAGTAA

M E A I W T A N I H D C I H G N I W P V

1570 1580 1590 1600 1610 1620
 1561 TGGGTCCAGC ATACCATAGC ATACACACA CGACATACA GCATACTAT GGTCAATTATA

M C A G Y H T I H H T T Y K H N Y G H Y

FIG. 8C

1621 1630 1640 1650 1660 1670 1680
 CCATATGCAT GCATTGCATG TTGGCTCTC TTACGGATCC TCTCTTACAA GAGATGACA
 T I H H D W H F G S L R D P L L E E D D
 1690 1700 1710 1720 1730 1740
 1681 ACAAGACAC CTCACGAAA GCAGACTGAG ATGCCCACT TCGCTTTTGT TCTTCTGTT
 N K D S F K K A E
 1750 1760 1770 1780 1790 1800
 1741 TGTCTTGTGT TGTGTGTGT CAAGATTCA GCCTTCTTG TCTTTTCT TCTTCTCT
 1810 1820 1830 1840 1850 1860
 1801 ATTATGTGT CTCTCTCAC CTTCACAT ATATTCTTAC AACATTTCG TGTCTAGTT
 1870 1880 1890
 1861 AAAACATGTA ATGTTTGTAT GATCTTTGA

FIG. 8D

1 MAADNAYLMQ FVDETSFYNR IVLSHLLPAN LWEPLPHFLQ TWLRNYLAGT
51 LLYFISGFLW CFYIYYLKIN VYLPKDAIPT IKAMRLQMFV AMKAMPWYTL
101 LPTVSESMIE RGWTKCFASI DEFGWILYFV YIAIYLVFVE FGIYWMHREL
151 HDIKPLYKYL HATHHIYNKQ NTLSPFAGLA FHPVDGILQA VPHVIALFIV
201 PIHFTTHIGL LFMEAIWTAN IHDCIHGNIW PVMGAGYHTI HHTTYKHNYG
251 HYTIWMDWMF GSLRDPLLEE DDNKDSFKKA E

FIG. 9

10 30 50
 GTTTGGTATTTATTGGATGCACAGAGAGCTTCATGACATTAAGCCTCTCTATAAGTATCT
 CAAACCATAAATAACCTACGTGTCTCTCGAAGTACTGTAATTCGGAGAGATATTCATAGA
 70 90 110
 CCATGCCACCCATCATATCTACAACAAGCAGAATACACTCTCTCCATTTGCCGGTAAGTG
 GGTACGGTGGGTAGTATAGATGTTGTTCTGCTTATGTGAGAGAGGTAAACGGCCATTCAC
 130 150 170
 TTTTCAGTTTGTCTTCTTTAGTTCTTGTAAGATTGGTAGCATTTAGTTTCTTACCAG
 AAAAGTCAAACAAGAAGAAATCAAGAACATTTTCTAACCATCGTAAATCAAAGAATGGTC
 190 210 230
 AAAAGACTTTGTGTCAGCAGCTGCTTGTAAGTCCAAATCACATTTGTCATTCTTATCCATAA
 TTTTCTGAAACAGTCGTGACGAACATGAGGTTTAGTGTAAGACGTAAGGAATAGGTATT
 250 270 290
 AGTAACCAGAAAGGCTAGAATTATATAAATGTCAGCTGCATTACTTCACATATGTCAGAG
 TCATTGGTCTTTCCGATCTTAATATATTTACAGTCGACGTAATGAAGTGTATACAGTCTC
 310 330 350
 AGACTTCTGACTTAACCAGAGTTTAGATCTTTGTGTTTCTCTTCTGGTCTCGGACTGATT
 TCTGAAGACTGAATTGGTCTCAAATCTAGAAACACAAAGAGAAGACCAGAGCCTGACTAA
 370 390 410
 GGAAATGACGAGAAGTTCTTTTATCTACTTCCCTGGAGTGTATCTTGGTTAATCCAAGGA
 CCTTTACTGCTCTTCAAGAAAATAGATGAAGGGACCTCACATAGAACCAATTAGGTTCTT
 430 450 470
 TGTGACATCTAAATATTACTTGTAACCTTACGTTTTTGTGTTACAGGGCTTGCAATTCA
 AACTGTAGATTTATAATGAACATTGAAGGAATGCAAAAACAAATGTCCCGAACGTAAGT
 490 510 530
 CCCAGTAGACGGGATACTTAAGGCTGTACCGCATGTGATAGCGCTGTTATAGTGCCAATT
 GGGTCATCTGCCCTATGAATTCGACATGGCGTACACTATCGCGACAATATCACGGTTAA
 550 570 590
 CATTTCACTCATATAGGTCTTTTGTTCATGGAAGCGATATGGACGGCGAACATCCAT
 GTAAAGTGTGAGTATATCCAGAAAACAAGTACCTTCGCTATACCTGCCGCTTGTAGGTA

FIG. 10A

610 630 650
 GACTGCATCCATGGCAACATCTGGCCAGTAATGGGTGCAGGATACCATACGATACACCAC
 CTGACGTAGGTACCGTTGTAGACCGGTCAATTACCCACGTCTATGGTATGCTATGTGGTG

670 690 710
 ACGACATACAAGCATAACTATGGTCATTATACCATATGGATGGATTGGATGTTTGGCTCT
 TGCTGTATGTTTCGTATTGATACAGTAATATGGTATACCTACCTAACCTACAAACCGAGA

730 750 770
 CTTAGGGATCCTCTCTTAGAAGAAGATGACAACAAAGACAGCTTCAAGAAAGCAGAGTGA
 GAATCCCTAGGAGAGAATCTTCTTCTACTGTTGTTTCTGTCTGAAGTTCTTTCGTCTCACT

790 810 830
 GAATGCCCACTTGGGTTTTGTTCTTCTGTTTTGTCTTGTGTTGTTGTTGTTCAAAGTTTC
 CTTACGGGTGAACCCAAAACAAGAAGACAAAACAGAACACAACAACAAGTTTCAAAG

850 870 890
 AGCCTTTCCTGTTCTTTTTCTTCTTCTTCTTATTTCATGTGTCTCTCTCAACCTTCCAAT
 TCGGAAAGAACAAGAAAAAGAAGAAGAATAAGTACACAGAGAGAGTTGGAAAGGTTA

910 930 950
 TATATTGTTACAAACATTTGCTGTCTAGTTTAAACATGTAAATGTTTGATGATCTTTGC
 ATATAACAATGTTTGTAACGACAGATCAAATTTGTACATTTACAACTACTAGAAACG

970 990 1010
 AAGACTCCATTTTGTGTTAAGGTAAACCTTGAATCTCATAGATTGTGCGATTGTTGGTATT
 TTCTGAGGTAAAAACAAATTCCATTTGGAACCTAGAGTATCTAACAGCTAACAACCATAA

1030 1050 1070
 TCCATTTTCAGGTACGGTTCTGTAGACTGTAGTCTTGCTGACCAGTCCGGCTTAACCACC
 AGGTAAAAGTCCATGCCAAGACATCTGACATCAGAACGACTGGTCAGGCCGAATTGGTGG

1090 1110 1130
 CCAAATTTCAAAGATCTCAcCAATCAAAATGCTGGCTGGCCCCAATATATAGATGGGCCA
 GGTTTAAAGTTTCTAGAGTgGTTAGTTTTACGACCGACCGGGGTATATATCTACCCGGT

1150 1170 1190
 GTTAATCCGTCTAGCTTTACTCTTTAGACCTACCTTAGACAGTTAGACACCTGCTAATTA
 CAATTAGGCAGATCGAAATGAGAAATCTGGATGGAATCTGTCAATCTGTGGACGATTAAT

FIG. 10B

1210	1230	1250
ATGAGTTTCCTTTTCTTGTTTCAGCAAGTTACCTGTGTTACTTGAGAGTTGAGTTAATGG TACTCAAAGGAAAAAGAACAAGTCGTTCAATGGACACAATGAACTCTCAACTCAATTACC		
1270	1290	1310
TAGTAAACGCAATTTAACCCTTATAAGTTTAAATCGTATTCAACGAATGACCCAGAGACTT ATCATTTCGCTTAAATTGGGAATATTCAAATTAGCATAAGTTGCTTACTGGGTCTCTGAA		
1330	1350	1370
TAAATAAATCCATCGTAACCCCTCCACTTCAAAATCTTTTAAAAAGTAGCAAATCATTT ATTTATTTAGGTAGCATTGGGAGGTGAAGTTTAAAGAAAAATTTTTCATCGTTTAGTAAA		
1390	1410	1430
AAATATTGTAAGTTTGCTTCATTTCGAAATTTGTAGCTACAGATCTCAAAGCTCCTCCTGTT TTTATAACATTCAAACGAAGTAAGCTTTAACATCGATGTCTAGAGTTTCGAGGAGGACAA		
1450	1470	1490
GGCCATATCTCTCTCTAACAACGCATAGTAACACTTGACCACAGTTTGACTTCTCGGCG CCGGTATAGAGAGAGATTGTTTTCGCTATCATTGTGAAGTGGTGTCAAAGTGAAGAGCCGC		
1510	1530	1550
GTTTCATGGCGCGACTATGGCAGATTATAATGATCAGATCGTCAATGAGACCTCTTTTT CAAAGTACCGCCGCTGATACCGTCTAATATTACTAGTCTAGCAGTTACTCTGGAGAAAAA M A A T M A D Y N D Q I V N E T S F Y		
1570	1590	1610
ACAACCGAATGGTTCTGAGTCACCTTTTGCCGgTGAATCTATGGGAACCTTTACCaCATT TGTTGGCTTaCCAAGACTCAGTGGAAAACGGCcACTTAGATACCCTTGGAATGGtGTAA N R M V L S H L L P V N L W E P L P H F		
1630	1650	1670
TCCTCCAGACATGGCTCCGGAACCTCGCCGGAACATACTCTACTTTCATCTCCGGCT AGGAGGTCTGTACCGAGGCCTTGATGGAGCGGCCTTTGTATGAGATGAAGTAGAGGCCGA L Q T W L R N Y L A G N I L Y F I S G F		
1690	1710	1730
TCCTCTGGTGCTTCTACATCTATTACCTTAAACTCAACGTTTACGTCCCCAAAGGTTACT AGGAGACCACGAAGATGTAGATAATGGAATTTGAGTTGCAAATGCAGGGGTTTCCAATGA L W C F Y I Y Y L K L N V Y V P K		

FIG. 10C

1750 1770 1790
 TTTTTC AATTTC GATG TTCT GTTTTG AAACCT TTTCTTTG TTGATT CCTTCG ATTGTATC
 AAAAAG TTAAAG CTACAAG ACAAACT TTTGGAAAG AAAACA ACTAAG GAAGCT AACATAG

1810 1830 1850
 GCCTGATAGATTGTGTTATACGTTAACCTTTTTTTCTTACTGTTACTTTTCAGTTCTTGTC
 CGGACTATCTAACACAATATGCAATTGGAAAAAAGAATGACAATGAAAGTCAAGAACAG

1870 1890 1910
 TTCTACTTCTCATTTAATTAGTTTTAAAGTTTAATATTTTTGGCTAATCCACATTTTTTA
 AAGATGAAGAGTAAATTAATCAAAATTTCAAATTATAAAAAACCGATTAGGTGTAAAAAAT

1930 1950 1970
 AGTTGAATCTTCCATGAAATTTGAGCTCAAAATATACCATGAAATTGAAATTTGTGGTTC
 TCAACTTAGAAGGTACTTTAAACTCGAGTTTTATATGGTACTTTAACTTTAAACACCAAG

1990 2010 2030
 TTAGTTCTATTTCTTGCTTGGTTTTCTTCTATTTTTGTGGTTAGAAATCCATTCCTACGAGA
 AATCAAGATAAAGAACGAACCAAAGAAGATAAAAAACCAATCTTAGGTAAGGATGCTCT
 E S I P T R

2050 2070 2090
 AAGGCAATGCTTTTGC AAATATACGTGGCAATGAAGGCTATGCCCTTGGTACACTCTTCTT
 TTCCGTTACGAAAACGTTTATATGCACCGTTACTTCCGATACGGAACCATGTGAGAAGAA
 K A M L L Q I Y V A M K A M P W Y T L L

2110 2130 2150
 CCAGCTGTCTCTGAGTATATGATCGAGCATGGTTGGACCAAATGTTACTCTACACTTGAC
 GGTTCGACAGAGACTCATATACTAGCTCGTACCAACCTGGTTTACAATGAGATGTGAAC TG
 P A V S E Y M I E H G W T K C Y S T L D

2170 2190 2210
 CATTTCAACTGGTTCCTCTGTTTCCTCTACATAGCTCTCTATCTTGTTTTAGTTGAGTTT
 GTAAAGTTGACCAAGGAGACAAAGGAGATGTATCGAGAGATAGAACAAAATCAACTCAAA
 H F N W F L C F L Y I A L Y L V L V E F

2230 2250 2270
 ATGATTTATTGGGTTCACAAAGAGCTTCATGACATTAAATTTCTCTATAAGCATCTCCAT
 TACTAAATAACCCAAGTGTTCGGAAGTACTGTAATTTAAAGAGATATTCGTAGAGGTA
 M I Y W V H K E L H D I K F L Y K H L H

FIG. 10D

2290	2310	2330
GCTACCCATCATATGTACAACAAGCAAAACACACTCTCTCCATTTGCCGGTATGTCAAAG		
CGATGGGTAGTATACATGTTGTTTCGTTTTGTGTGAGAGAGGTAAACGGCCATACAGTTTC		
A T H H M Y N K Q N T L S P F A		
2350	2370	2390
CTATATGTTCTCAATCTAAATTCAAGAGCTTGTATCAATGGTGACTTCTTTACTTGATGT		
GATATACAAGAGTTAGATTTAAGTTCTCGAACATAGTTACCACTGAAGAAATGAACTACA		
2410	2430	2450
TTTTTCGGGTTTTTCAGGGCTCGCATTCCATCCGCTGGACGGGATACTTCAGGCCTATACCGC		
AAAAGCCCCAAAAGTCCCGAGCGTAAGGTAGGCGACCTGCCCTATGAAGTCCGATATGGCG		
G L A F H P L D G I L Q A I P H		
2470	2490	2510
ACGTGATAGCGCTGTTTATAGTGCCGATTTCATCTCATAACACATCTGAGTCTTTTGTITTT		
TGCCTATCGCGACAAATATCACGGCTAAGTAGAGTATTGTGTAGACTCAGAAAACAAAA		
V I A L F I V P I H L I T H L S L L F L		
2530	2550	2570
TGGAAGGGATATGGACAGCAAGCATCCATGATTGCATACATGGtAACATCTGGCCTATAA		
ACCTTCCCTATACCTGTCGTTTCGTAGCTACTAACGTATGTACCaTTGTAGACCCGATATT		
E G I W T A S I H D C I H G N I W P I M		
2590	2610	2630
TGGGTGCAGGATACCATACCATAACCATAACAACATAACAAGCATAACTATGGTCATTATa		
ACCCACGTCCTATGGTATGGTATGTGGTATGTTGTATGTTTCGTATTGATACCAGTAATAt		
G A G Y H T I H H T T Y K H N Y G H Y T		
2650	2670	2690
CCATATGGATGGaCTGGATGTTTGGCTCTCTTATGGTTCCTTTAGCAGAAAAAGACAGTT		
GGTATACCTACCtGACCTACAAACCGAGAGAATACCAAGGAAATCGTCTTTTTCTGTCAA		
I W M D W M F G S L M V P L A E K D S F		
2710	2730	2750
TCAAGGAGAAAGAAAAGTGAGAATGTTCAATGCTCACATGTATTCTTTCATATGTTGCTCT		
AGTTCTCTTTCTTTTCACTCTTACAAGTTACGAGTGACATAAGAAGTATAACAACGAGA		
K E K E K *		
2770	2790	2810
TCTCGTGACTCTTATTAAAACCTTTCTAATCACTTTGGTGGAATTAAAAACATGACTGCA		
AGAGCACTGAGAATAATTTTGAAAGATTAGTGAAACCACCTTAATTTTGTACTGACGT		

FIG. 10E

2830 2850 2870
TAATTTGATGCAAAGTTTCAGACTTTTATTGCTAAAAATCTCTGATGATTATTAACCTCA
ATTAAACTACGTTTCAAAGTCTGAAAATAACGATTTTTAGAGACTACTAATAATTGGAGT

2890 2910
ATTATATAATTGcTGGATGAAGAGTTCAAATTTGGACTAAATCTG
TAATATATTAACgACCTACTTCTCAAGTTTAAACCTGATTTAGAC

FIG. 10F

1 maatmadynd qivnetsfyn rmvlshllpv nlweplphfl qtwlrnylag
51 nilyfisgfl wcfyiyylkl nvyvpkesip trkamllqiy vamkampwyt
101 llpavseymi ehgwtkcyst ldhfnwflcf lyialylvlv efmiywvhke
151 lhdikflykh lhathhmynk qntlspfagl afhpldgilq aiphvialfi
201 vpihlithls llflegiwta sihdcihgni wpimgagyht ihhttykhny
251 ghytiwmdwm fgslmvplae kdsfkekek

FIG. 11